

Combined Digital Preservation Resources for WYPR:

Survey Report, Next Steps Plan, and Digital Preservation Policy

Margaret McCready, MLIS Candidate
University of Maryland, College Park
December 2nd, 2018

Table of Contents

Executive Summary:	3
Survey Report of Digital Content and Management Practices at WYPR	4
Survey Summary:	4
WYPR's Mission and Work:	5
Scope of WYPR's Holdings:	5
Digital:	5
Analog:	6
Current Management Practices for Digital Holdings:	6
Digital:	6
Analog:	8
Staff Perceptions of the State of Digital Content:	9
Gaps in WYPR's Collection and Potential Collection Interests:	9
WYPR Staff Resources and Abilities:	10
Institutional Next Steps Plan for WYPR	10
Introduction:	10
Structure of the Plan:	11
Summary of Recommendations:	11
Storage and Geographic Location:	12
File Fixity and Data Integrity:	13
Information Security:	14
Metadata:	15
File Formats:	16
Digital Preservation Policy for WYPR and Midday	16
Introduction	16
Scope and Selection of Materials for Preservation	17
Roles and Responsibilities	17
Records Life Cycle	18
Strategies and Preservation Actions	18
Documentation and Standardized Practice	18
Storage and Maintenance	18
File Formats and Standards	19

Executive Summary:

The contents of this report were developed in collaboration between Margaret McCready (MLIS Candidate, UMD), Rob Sivak (Senior Producer, *Midday*) and Tom Hall (Host, *Midday*). Margaret McCready produced all written materials presented in this report. The project initially spanned between October 12th, 2018 and December 2nd, 2018.

An initial survey of the station revealed several concerns relating to current management practices, that have potentially catastrophic consequences if not addressed quickly. At present, all files are stored at the station and only one copy exists of all production files and audio recordings. In the event of a disaster, there is a significant chance of complete loss of all files.

Additionally, master recordings of broadcasts are stored only on CD-Rs, which specifically face high risk of loss due to theft, disaster, or natural degradation of the medium. This is not a stable or reliable long term storage solution.

Current description and storage practices for production files and audio are not standardized, and as a result, records are dispersed across different storage media, and are not easily locatable. This reduces efficiency and impairs the ability of WYPR staff to respond promptly to requests from the community or the Federal Communications Commission, or to locate materials to be used for future broadcasts.

The next steps plan details necessary actions to improve the stability and accessibility of WYPR's files. The plan recommends that WYPR begin to store at minimum **two** copies of all production files and audio, on both an external hard drive and a third party cloud storage service. This will improve security of files while also mitigating risks of loss in the event of a disaster. CD-Rs must be phased out of use. A comprehensive inventory of all files will ensure that staff have an accurate understanding of their collections, and allow staff to periodically check that all files are accounted for. The inventory will also provide information about the content of broadcasts and production files to improve efficiency of locating and accessing files. Finally, the Producer's Manual should be updated to provide accurate and up-to-date guidance on best practices for managing files at WYPR.

The Digital Preservation Policy provides overarching guidance for what materials should be saved, responsibilities of staff to preserve files, the life cycle of materials, and strategies for preserving materials. The policy reiterates the next steps plan, stating that it is the responsibility of producers to maintain production files and audio recordings in a series of nested folders on the external hard drive, and transfer files to a third party cloud service. Producers must also maintain an accurate inventory of all files. Interns are not to have access to master copies of files on the external hard drive, for security purposes.

Survey Report of Digital Content and Management Practices at WYPR

Survey Summary:

This survey report of digital management practices at WYPR was produced by Maggie McCready, a graduate student at the University of Maryland, College Park, following an interview with Tom Hall and Rob Sivak on October 12th, 2018. The results of this survey were based upon the management practices of Senior Producer Rob Sivak of *Midday*, in addition to information provided by Host Tom Hall.

WYPR's collections of digital content primarily include digital audio recordings of its broadcasts, as well as production files or "script packs" used to produce those broadcasts.

Current storage methods for these materials include audio stored on CDs which is copied to an external hard drive annually; and "script packs" are typically Microsoft Word documents stored in a series of nested folders organized by date in the station's shared server, though additional copies are also stored on the external hard drive. All materials are stored at the WYPR station.

While current practices are acceptable, there are several potential risks that threaten the stability and longevity of WYPR's digital content.

At present, all of WYPR's materials are physically stored (including the server's physical location) at the station. Without more geographically diverse storage practices, all of WYPR's content and files could be lost in the event of a disaster.

Additionally, there is currently no fixity checking or security/access restrictions being performed to ensure the authenticity and integrity of files stored at WYPR. Without regular fixity checks, stored digital content may be lost due to bit rot. (A further explanation of fixity checking is provided below in the Current Management Practices for Digital Holdings and Risks section of this report.) Without security or access restrictions, files could be accidentally or purposefully deleted or altered.

Current management practices at the station are quite individualized, with no current, standardized policy for preservation that is universally followed. Currently, producers are responsible for managing all the digital content associated with their respective shows, though there have been instances of producers or hosts leaving the station and taking their recordings and work files with them, leaving significant gaps in WYPR's collections. Implementation of accountability and policy changes could address these issues.

Additional information regarding these topics, including analog media, metadata practices, and file formats, are provided in the report below.

WYPR's Mission and Work:

"The mission of Your Public Radio is to inform, connect and even challenge the listeners we serve in the metropolitan Baltimore area and the State of Maryland by broadcasting programs of intellectual integrity and cultural merit so as to provide an unbiased perspective of the events of today and to enrich the minds and spirits of our audience." WYPR is a public radio station, which is a local branch of NPR.

As such, WYPR serves as a news and communication resource for the greater Baltimore region and the State of Maryland, and generates several local programs, including *Midday*, *On the Record*, *Gil Sandler's Baltimore Stories*, *Out of the Blocks*, and others.

These programs serve to document local experiences and concerns of/for residents of the State of Maryland, and could thus be considered records of significant historical, research, or community value.

Scope of WYPR's Holdings:

Digital:

At present, WYPR maintains digital recordings of each local program it produces, in addition to "script packs", which include production notes, scripts, and other promotional content used to create these programs.

- Digital recordings of shows are stored on CD-RW (Compact Disk), and are considered the "Master Files" of the audio, which are created by the producer of the show by exporting 4 uncompressed .WAV files from Cool Aud/Adobe Audition. However, there are also .MP3 versions of these files on the disk. The audio stored on the disk includes the actual show, but also a promotional "billboard" recording. There are recordings dating back to 2006 for shows including *Midday* stored on these disks, and approximately 6 linear feet of material.
- Script packs for the shows are typically Microsoft Word documents. Again, this content largely includes the "raw material" associated with creating each show, as well as promotional material, dates and titles of each show, written scripts, and billboards. This content is largely "Born Digital".
- WYPR also utilizes a number of digital tools in the production of their shows, such as a shared Google Calendar, with lists of dates and titles of shows being produced; "Pleats", a modified Microsoft Access database that has all information about shows,

demographic data about guests, and more. This database is stored on the shared drive/server at the station.

Analog:

While the scope of this project is to address digital preservation issues, it is also prudent to draw attention to the type and scope of WYPR's analog materials.

- In the storage closet by the recording studios, approximately 10-15 linear feet of 12" and 6" open reel magnetic tapes were found, dating to the late 1990s and early 2000s. (However, some of this material appears quite older.) This material was likely produced when the station was WJHU (under Johns Hopkins). Due to unideal storage conditions, this material is classified at **high risk for potential loss**, and suggestions for this material will be provided in following documents.
- Additionally, several DAT tapes were identified in the collection. These were labeled with program titles and dates. However, their extent is unknown.
- There are other as of yet unidentified magnetic tape storage media kept in this closet, however, their extent and content is unknown.

Current Management Practices for Digital Holdings:

At present, the author is primarily aware of the management practices specifically undertaken by Rob Sivak, Senior Producer of *Midday* at WYPR. Management practices of other producers at WYPR is presently unknown.

However, current digital management practices do meet some of the requirements for the first level of digital preservation as outlined by the [National Digital Stewardship Alliance's \(NDSA\) Levels of Digital Preservation Model](#). Specifically, WYPR meets the basic requirements for the first level of the "Storage and Geographic Location" category, the "Metadata" category, and the "File Format" category. Though at present, there are no steps being taken to guarantee "File Fixity and Data Integrity", or "Information Security". This will be addressed in greater detail below.

Digital:

Storage and Geographic Location Practices:

- WYPR maintains two copies of each program's digital audio recordings on two different types of storage media. The audio is initially burned to a Compact Disk(CD), which is then kept in a binder on a shelf at the station. This is the only copy of the audio.

- Further, the audio associated with WYPR's shows are also used to produce podcasts, which are hosted on WYPR's website using AudioStack, a platform for advertising and generating revenue. Older shows can be accessed through this platform. However, there is not much control over the files stored this way.
- The Script Packs are stored in a series of nested folders on the shared server, organized by date of the show, and these files are periodically moved off the server onto an external hard drive.
- The digital tool Google Calendar is not currently backed up or stored physically by WYPR. The "Pleats" database is kept on the shared server, but it is unknown if additional copies exist.
- **Potential Risk:**
 - WYPR's files are stored either on a shared server at the station, a CD stored at the station, or an external hard drive located at the station. There are no copies stored in a different geographic location, and in the case of a fire or other disaster at the station, all copies of the digital audio recordings and script packs would be lost.
 - Currently, WYPR does not monitor the condition of their storage media for obsolescence.

File Fixity and Data Integrity:

File fixity refers to the bits of a file. "Fixity information offers evidence that one set of bits is identical to another." (NDSA, 2014,1.) With fixity information, a user can identify if a file has become corrupted, or altered in an unauthorized way, by checking to see if two copies of the same file are identical. Fixity information serves as the "fingerprint" of a file.

For more information about File Fixity: [What is File Fixity, and When Should I Be Checking It?](#)

At present, WYPR does not perform any kind of fixity checking.

- **Potential Risk:**
 - Without monitoring the state of files periodically or during file transfers, there is a risk that files could be degrading over time or the authenticity of files could be at risk.

Information Security:

The majority of digital files, including audio and script packs are stored on the station's shared server, CDs, or an external hard drive that is kept on a desk. Currently, there are no controls on who has authorization to read, write, edit, or delete files.

- **Potential Risk:**

- By storing copies of files on a shared server that anyone at the station has access to, there is the potential for overwriting or accidental deletion of records. External hard drive and CDs are easily accessible, and could be either taken or destroyed.

Metadata:

Metadata practices at WYPR are not standardized, with no standard formats for file naming conventions. However, script packs and the associated “raw material” created for each show are well organized, and are kept in nested folders on the shared server arranged by the date of the show. There is considerable descriptive metadata generated for each file, including the title or purpose of a document, which show it was produced for, and the date. CDs have written information on the front regarding the show, date aired, and what segments were a part of that show. It is unknown whether or not there is a complete inventory of all files and their storage locations.

- **Potential Risks:**

- Current descriptive metadata practices are sound, though without an inventory of files and adequate administrative metadata, it is easy for files to get lost. This is especially a concern if the employee who managed those files left the station and left no written information about what they did with the files.

File Formats:

The majority of files generated at WYPR are either .DOC, .WAV, or .MP3 files, though there are likely Microsoft Excel or .JPG files as well. The usage of these file formats, particularly the .WAV and .MP3 formats for recorded audio, appear to be standardized practice. The use of such limited formats improves the capacity to manage those files.

- **Potential Risks:**

- These file formats are relatively common and do not face substantial risk of becoming obsolete or inaccessible. Though an inventory of file formats used at the station would confirm if there are any risks.

Analog:

Presently, all CDs containing digital audio from shows are kept in CD binders on a shelf in an office at the station. Older analog materials, such as 12” and 6” open reel magnetic tapes are stored in a closet outside the recording studios on the first floor of the station. Some of the open reel tapes are not stored properly in a case, often with the tapes unraveled and hanging. The DAT tapes are stored in boxes on the shelves in this same closet.

- **Potential Risks:**
 - There is extreme risk of loss of the 12” and 6” open reel magnetic tapes due to inadequate storage conditions. Action needs to be taken to preserve this material. The CDs are in relatively stable conditions, though the quality of their storage containers are questionable. The DAT tapes are susceptible to abrasive dust in their current environment.

Staff Perceptions of the State of Digital Content:

Upon speaking to *Midday* host, Tom Hall and senior producer Rob Sivak, current management practices are acceptable and actually exceed initial expectations of the author when compared to the practices of other, similarly sized cultural institutions.

However, both Hall and Sivak expressed desire to improve practices and indicated that practices at the station are not standardized, but left to the individual discretion of the producers for each show. There is guidance for how to record and store audio from shows from an outdated version of a “Producer’s Guide”, however this is not sufficient alone to instruct new people or interns on how to preserve audio, often verbal instruction is additionally required.

The main purpose for preserving digital content at WYPR is for legal and rebroadcast purposes. For accountability and legal purposes, WYPR must maintain recordings of all broadcasts in the event that a complaint is filed the Federal Communications Commission (FCC) against WYPR. Additionally, at the end of each calendar year, WYPR chooses 5 shows to rebroadcast, and ease of accessibility to this information would help increase efficiency.

Gaps in WYPR’s Collection and Potential Collection Interests:

Both Sivak and Hall expressed that current collection practices are sufficient, with no desire to begin collecting or saving additional material. This is outside of the scope of WYPR’s mission.

However, there were concerns regarding gaps in recordings from previous hosts of the *Midday* radio program, that is missing from WYPR. This is not a particularly pressing concern, but could potentially be addressed.

Additionally, Sivak indicated that many of the shows at the WYPR station are preserved or stored differently and in disparate places. There is interest in perhaps developing a more universal or standardized method for storing production files and audio recordings, particularly if this could lessen the workload or take less time for the already busy producers at WYPR.

WYPR Staff Resources and Abilities:

With regard to the staff and WYPR's ability to dedicate its efforts to sustaining their digital content, both Hall and Sivak indicated that time and funds are an issue. Many of the shows at the station are considered understaffed, (WYPR's *Midday* has 2 full time producers, compared to 9 producers that worked on the Diane Rehm Show hosted by WAMU in DC.) Meaning that time is tight, and staff at WYPR could not reasonably commit additional work hours to digital preservation. However, WYPR also takes on a few interns throughout the year, and some of this work could be assigned in addition to their regular duties.

Funding is a concern as a community radio station, however moderate changes in spending to improve WYPR's management of digital content are possible. Sivak considered the possibility of investing in a cloud-based storage service to help diversify geographic locations of file storage and mitigate the impact of disaster related loss. Additionally, Sivak and Hall considered the possibility of collaborating with other local branches of NPR to store additional copies of digital files (diversifying geographic risk.)

With regard to the analog materials described in this survey report, it is completely out of scope of WYPR's mission to serve as a repository or to dedicate time to preservation of these materials. Considering that the majority of the open reel magnetic tapes were originally created by WJHU prior to WYPR's founding in 2002, donation of these materials to a proper repository (Johns Hopkins Special Collections, or University of Maryland Special Collections) was discussed as a possibility.

Despite financial and time barriers, minor shifts in current practices and investments in better storage options seem to be actionable responses.

Institutional Next Steps Plan for WYPR

Introduction:

For clarity, the recommendations provided in this plan are based on the [National Digital Stewardship Alliance's Levels of Digital Preservation](#) (NDSA LoDP). This plan attempts to improve WYPR's "level" of digital preservation practices in areas of Storage and Geographic Location; File Fixity and Data Integrity; Information Security; Metadata; and File Formats. These 5 elements are represented in a table with 4 progressive levels of quality of practice. The goal of this project is to improve WYPR's practices from Level 0/1 to Level 2.

However, the recommendations provided in this plan were made in recognition that WYPR is not a cultural heritage repository (which the NDSA Levels of Digital Preservation are intended for), and as such its needs are different. The intention of this plan is to improve current management practices to a reasonable level without overburdening producers and WYPR staff.

Structure of the Plan:

This plan will break down suggestions for next steps into the aforementioned 5 categories, which are further broken down into Short Term, Mid Term, and Long Term goals, based upon the level of effort required to implement, or the urgency of the action. At the end of each recommendation, the level of the action will be rated according to the NDSA's levels (e.g. (Level 1-4.))

While Long Term actions may require additional effort or take longer to complete, these are still necessary to ensure sustained improvement of practice for managing digital content at WYPR.

Summary of Recommendations:

In brief, the plan identifies two specific steps that would significantly improve WYPR's digital preservation practices and effectiveness without a significant increase in workload or staff time.

First, the plan recommends **investment in a third party cloud storage platform**, such as Carbonite or Dropbox, to begin storing a second copy of all digital files, diversify geographic storage of files and mitigate disaster risk, and introduce basic fixity checking of all files. These services require little effort to maintain and costs of services are offered on a sliding scale.

Second, the plan recommends developing a **comprehensive inventory of all files** (script packs, audio recordings) in Excel spreadsheet. This inventory should include documentation of expected file size and amount, descriptive information about files (file name or content), administrative information (storage location, date created, file format, date transferred or copied), and who implemented those changes.

This is a multi use tool that provides a clear list of all production files in one easily accessible place, allows basic fixity checking, introduces documentation of how files are managed, and prevents confusion or loss of information that may occur when staff retire or leave the station. While the spreadsheet may require persistent upkeep, this is a task that may be delegated to an intern, which may be checked for accuracy periodically by WYPR staff.

Additional recommendations include:

- Updating the Producer's Manual to provide up to date explanations of the process of recording and storing digital audio from broadcasts. This will increase efficiency and reduce staff time devoted to verbally training interns or new staff.
 - Include guidance for file naming conventions, descriptive metadata, acceptable file formats, and the process of transferring files.
- Locking or otherwise restricting access to files stored on the external hard drive for security purposes.

Finally, it is recommended that WYPR staff discuss the state of the WJHU open reel magnetic tapes stored in the first floor closet. These tapes are in poor condition and are outside of the scope of WYPR, and it is recommended that they be donated to either Johns Hopkins, the [American Archive of Public Broadcasting](#), or another similar repository for archival material.

Storage and Geographic Location:

Currently, WYPR staff maintain one to two copies of their digital files, with script packs kept both on the station's internal server and an external hard drive, while digital audio of the shows are only stored on CD-Rs. This adheres to the NDSA's Level 1 recommendations for storage, but this could be improved as high as Level 3.

In order to achieve higher quality of care for digital content, it is recommended that WYPR begin to store at least 2 copies of all files in different types of storage media, which are then kept in different geographic locations that face different disaster threats. Additionally, a written log should be kept of all types of storage solutions utilized and how they can be accessed for standardization and efficiency purposes. Specific recommendations will be provided below:

Short Term:

- Create a list of all storage media currently utilized, and provide directions on how to access digital content stored on that device. (Level 2).
- Get at least one copy of all files stored in a different geographic location, either at someone's home, in the cloud, or at another NPR station. (Level 2).
- Create a second or third copy of all script packs and digital audio content and store it on a third type of storage media. (Level 2).
- Discuss possible options for diversifying geographic risk (storing files offsite), potential partnerships with local NPR stations for storing each other's files, consider costs.

Mid Term:

- Assemble and arrange all files pertaining to *Midday*, including those created by past hosts and producers, in a hierarchical series of folders on the shared drive to ensure that

files relating to the show are easily accessible. Mimic this structure for files stored on other storage media (e.g. the external hard drive, or cloud storage.)

- To diversify geographic storage and disaster risk, invest in a cloud storage service (Carbonite, Drop Box, Google Drive) or an additional external hard drive. Copies of all files should be transferred to these locations periodically at a designated time (monthly, quarterly, annually). (Level 3).
 - Cloud storage options are automatically geographically dispersed, have built in access controls for security of data, and have unlimited storage, but typically charge monthly and there is slightly less control over data. The typical monthly charge can be anywhere between 20 to 50 dollars a month depending on services. This would be an efficient means of “killing four birds with one stone”, as cloud storage services address concerns of storage space, disaster risk, security, and to an extent file fixity.

Long Term:

- Address issues with inconsistent or out of date directions for digital file management in the Producer’s Manual. Develop a written procedure for how to store digital audio and script packs, including a requirement to maintain 2 copies on 2 different storage media, and outline when files need to be copied or transferred to different storage media. Outline what is and what is not acceptable as a storage device (e.g. no floppy disks.) This step is crucial to ensure consistency and adherence to standardized procedures, which ultimately will streamline the management of and access to WYPR’s files. (Level 2-4).
 - This process of creating procedure will be part of the next step of the author’s project.
- Begin to consider alternatives to relying upon CDs as storage method for digital audio, as these have a high failure rate beyond 7 years. External hard drives, server storage, or cloud storage are all more reliable alternatives.
- Monitor storage media for degradation or obsolescence. (Level 3).

File Fixity and Data Integrity:

WYPR does not perform or maintain any fixity checking on its digital content. Fixity checking will ensure the long term preservation and integrity of files by identifying issues that arise from transferring or copying files. Additionally, “If checks against fixity information for a set of objects begin fail at high rates, it can be an indication of media failure.” ([What is Fixity, and When Should I be Checking it?, 2014, 2](#)). This is especially important for the audio recordings stored on CDs, which is a highly failure prone storage media, to make sure that files are not lost to bit rot.

A simple response to this problem would be to create an inventory of all script packs and digital audio recordings, which includes expected file count and size, file format, when a file was created or transferred, and who took that action. This will get WYPR to at least Level 1 of the LoDP, and is less time consuming than performing individual fixity checking on each file. However, this inventory will need to be kept up-to-date, and an audit of all files should be performed at least annually to ensure that “everything’s where it’s supposed to be”.

Short Term:

- Perform a basic fixity check by creating an excel spreadsheet inventory of all files, including script packs and digital audio recordings, writing down file size of each nested folder and amount of files in each folder. Additional information documenting when files are transferred to different storage media can be kept in this spreadsheet. (Level 1).

Mid Term:

- Begin using a system such as BagIt to gather all files being transferred to the external hard drive or cloud storage service, which automatically generates checksum (fixity information) for all files contained within as a simple .TXT file. (Level 1-2).
- Utilize third party cloud storage service, such as Carbonite, to run fixity checks of stored data. Most cloud storage services offer this service, though the frequency and detail of these checks varies from service to service. Ultimately, this can save time for production staff at WYPR. (Level 3).

Long Term:

- Check fixity of all files on an annual basis by referencing the file inventory to check if file size and amount are accurately depicted, or by comparing original fixity information (check sums) to newly generated check sums. Basically, this is just to check that everything you think is there, is actually there. (Level 3).

Information Security:

Current security practices do not meet any of the requirements for the NDSA Levels of Digital Preservation, as files are stored either on a station wide server, stored in an external hard drive which is kept on a desk, or CDs which are on a shelf in an open office. These files are easily accessible to anyone in the station, and there are no formal instructions on who has read, write, move, and delete authorization, nor logs of what actions have been taken with files.

The most immediately actionable response is also the simplest: lock up external hard drives or offices when they are not in use. Additionally, the host and producers of *Midday* should discuss

access restrictions for their content, making this known to other staff at the WYPR station. Doing both of these things will easily bring information security practices to Level 2.

Short Term:

- Lock external harddrive in a drawer or safe when not in use to prevent tampering or theft of files. Or, lock office when not at the station. (Level 1).
- Document access restrictions for content. Create a written document that states who can read, who can write, and who can delete or move files. Make this known to other staff at the station. (Level 1/2).

Mid Term:

- Maintain an excel sheet of who copied, edited, or deleted files and when, particularly during the transfer of files from one storage media to another. This record of files can also serve as a useful inventory of all digital content and can be used to perform basic fixity checking. This dual use tool is especially important. (Level 3).

Long Term:

- Check logs annually to ensure that files have not been altered. (Level 4).

Metadata:

Metadata for current files are limited, with primarily administrative data being generated that documents when the files were created. Much administrative metadata such as file type, date, and what program was used to create it is automatically generated and kept by the operating system. The focus will be to improve descriptive metadata practices to make locating and identifying content of past shows easier for producers, particularly when identifying programs for rebroadcast at the end of the year.

Short Term:

- Create excel spreadsheet inventory of all script packs and audio recordings created for *Midday* to establish a sense of WYPR's holdings. (Level 1).
- Create standardized file naming conventions that clearly describe the content of the script pack, or audio recording. Consider the important elements of what is being described-- date of the show? Subjects/Topics? Guests? (Level 3).

Mid Term:

- Document administrative metadata in excel spreadsheet inventory, including date created, who created, file format, and documentation of file transfers. (Level 2).

Long Term:

- Maintain file inventory and adhere to established file naming conventions.

File Formats:

There are currently several different types of files utilized by staff at WYPR, including .DOC, .WAV, and .MP3 files, though the producer's manual provides some guidance on the required formats for recording digital audio from broadcasts. These are fairly common file formats and do not face significant risk of obsolescence. This currently meets Level 1 standards, but additional standardization and documentation of acceptable file formats is necessary.

Short Term:

- Create a list of all file formats currently in use at WYPR. (Level 2).
- Create a limited, standardized list of acceptable file formats. (e.g. don't use .TXT files for script packs, or don't use .WMA for audio recordings.) (Level 1).

Long Term:

- Include documentation of acceptable file formats in updated iteration of the Producer's Manual. (Level 1).

Digital Preservation Policy for WYPR and *Midday*

Introduction

The Your Public Radio (WYPR) station in Baltimore, Maryland serves as a news and communication resource for the greater Baltimore region and the State of Maryland. WYPR is a public radio station, and a local branch of NPR. WYPR generates several local programs, including *Midday*, *On the Record*, *Gil Sandler's Baltimore Stories*, *Out of the Blocks*, and others.

These programs serve to document local experiences and concerns of/for residents of the State of Maryland, and could thus be considered records of significant historical, research, or community value.

Considering that the majority of all production files, audio, and administrative files at WYPR are born-digital, implementing a sound digital preservation policy based upon archival best practices emerges as a necessity to ensure the longevity of these locally significant materials.

Preservation of digital records is also necessary for accountability and legal purposes, to enable WYPR to respond to requests from researchers, members of the community, and the Federal Communications Commission.

This policy was specifically drafted in recognition of WYPR's needs, capabilities, and mission. As such, the practices put forth in this plan do not require specialized training, or drastic increases in staff time to implement, and can be adjusted if issues arise.

Scope and Selection of Materials for Preservation

The scope and purpose of this policy is to address files generated during the production of radio broadcasts, as well as the broadcasts themselves. Administrative files that document the overall management of the station, as well as outreach initiatives are not included in this policy, though this may be amended in the future.

Materials that should be saved:

All materials generated during the production of radio broadcasts, or that were referenced or influenced the structure or content of the broadcast. This includes:

- Meeting notes
- Production notes and planning documents
- Emails or correspondence between producers, hosts, or guests
- Script packs/written scripts
- Promotional materials (i.e. billboards)
- Paper work
- Spreadsheets
- Photographs or images

In documenting the broadcast itself:

- Master recording of broadcast
- Recordings of any recorded promotional spots (i.e. billboards)
- Any post-production files

Roles and Responsibilities

Preservation of production files and audio recordings falls under the responsibility of production staff for *Midday*, with the majority of preservation tasks undertaken by the senior producer.

Station interns may also contribute to the maintenance of spreadsheet inventories, and limited access to production files on the server, but should not have access to “master” copies of files stored on external hard drives or cloud services.

Records Life Cycle

It is the responsibility of producers and hosts to collaboratively make a decision regarding the long term storage of their records.

Strategies and Preservation Actions

Specific strategies for preserving digital materials differ from format to format, and there is no “one size fits all” solution for digital preservation. Additionally, it is best to keep in mind that one is never “finished” with digital preservation, it is an ongoing process.

The National Digital Stewardship Alliance’s [Levels of Digital Preservation](#) provides documentation of best practices for management of digital content. These recommendations are intended for cultural heritage repositories; however, as a small, public radio station, WYPR is only capable of sustainably performing Level 2 preservation actions.

Documentation and Standardized Practice

Accurate, and up-to-date information regarding practices should be maintained in the Producer’s Manual. Additional documentation includes descriptions of the size and extent of the *Midday* archive, what is missing and why, and how files from the archive can be accessed.

Accurate inventories of all files in the *Midday* archive should be maintained monthly. This should be maintained at the daily folder level, individual files do not need to be included on the inventory. Documentation of this process will be provided in the Producer’s Manual.

- The inventories and corresponding files will be checked against each other bi-annually to ensure that all files are present. This will ensure fixity (authenticity) of all files. If files are lost or corrupted, additional copies of files may be used to replace them.

Storage and Maintenance

- WYPR will maintain two copies of all digital files in different storage media and different geographic locations. Maintenance of redundant copies will allow for backup in the event of loss or disaster. This will include storage on an external hard drive and a cloud storage service.
 - Script packs will be transferred from the station server to an external hard drive **daily**, after the broadcast for that day has aired.

- WYPR will begin to **phase out use of CD-Rs** as storage media. This medium is prone to loss and data stored on these disks is less efficient to retrieve.
 - All digital audio currently stored on CD-Rs must be transferred to both the external hard drive and cloud storage platform.
- For security purposes, the external hard drive must be placed in a locked drawer when not in use.
- WYPR will adhere to a standardized file naming structure for script packs and audio. Guidelines will be provided in the Producer's Manual. For further reference on this matter:
 - ["5 Tips For What NOT To Do When Creating A File Naming Structure"](#)
- All files will be stored alphabetically by program title, thereunder chronologically by year, thereunder by individual date. A master folder is created for each program (e.g. *Midday*), which contains **all** files pertaining to the show. Further explanation is provided in the Producer's Manual.

File Formats and Standards

WYPR will maintain a list of acceptable file formats in its Producer's Manual, and limit the variety of formats used in order to ensure ease of access and preservation.¹

Acceptable File Formats for Text Documents:

Microsoft Word .doc Microsoft PowerPoint .ppt
 Microsoft Excel .xls PDF .pdf

Acceptable File Formats for Images:

JPEG .jpg
 PNG .png
 TIFF .tif, .tiff
 Photoshop .psd
 JPEG2000 .jp2

Acceptable File Formats for Audio:

MPEG audio .mp3 Wave .wav

¹ Documentation of preferred digital file formats was provided by the [Boston University Libraries Digital Preservation Policy](#).